

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
22 March 2001 (22.03.2001)

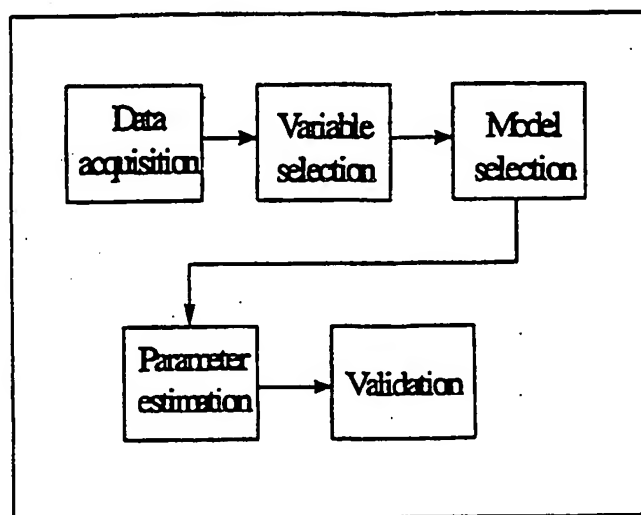
PCT

(10) International Publication Number
WO 01/20512 A2

- (51) International Patent Classification⁷: G06F 17/60 (74) Agents: CHASKIN, Jay, L. et al.; General Electric Company, 3135 Easton Turnpike W3C, Fairfield, CT 06431 (US).
- (21) International Application Number: PCT/US00/24414
- (22) International Filing Date:
6 September 2000 (06.09.2000) (81) Designated States (*national*): AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW.
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
09/396,599 15 September 1999 (15.09.1999) US
- (71) Applicant: GENERAL ELECTRIC COMPANY
[US/US]; 1 River Road, Schenectady, NY 12345 (US).
- (72) Inventors: CHEN, Yu-To; 1223 Carlyle Drive, Niskayuna, NY 12309 (US). BONISSONE, Piero, Patrone; 1065 Avon Road, Schenectady, NY 12308 (US). TRENCH, Margaret, Stewart; 116 Crane Street, Scotia, NY 12302 (US). DONOGHUE, Jeremiah, Francis; 1 Blue Spruce Lane, Ballston Lake, NY 12019 (US).
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:
— Without international search report and to be republished upon receipt of that report.

[Continued on next page]

(54) Title: METHOD FOR MODELING MARKET RESPONSE RATES



(57) Abstract: A method for modeling marketing response rates. The method is used to evaluate and filter large contact lists with the aim of accomplishing two goals. The method of the present invention uses an internal experience database and an external demographic database, coupled with variable screening schemes and non-parametric modeling techniques. The method comprises a number of steps. First, a data acquisition step associates descriptor variables with prospects. Second, a variable selection step identifies the descriptor variables in order to identify prospects most likely to respond to the direct mailing. Third, a model selection step examines and assesses a number of competitive algorithms, and selects the algorithm that will best predict the response rate. Fourth, a parameter estimation step ensures the best fit of data once an algorithm is chosen. Finally, a validation step ensures the robustness of the modeling process.

WO 01/20512 A2



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHOD FOR MODELING MARKET RESPONSE RATES

FIELD OF THE INVENTION

The present invention generally relates to direct marketing, and more specifically to modeling market response rates to direct solicitations.

BACKGROUND OF THE INVENTION

Direct marketing usually involves directly contacting persons or entities, such as for example by mail, with a specific message or solicitation. The persons to be contacted are usually identified by a mailing list. Today's direct marketer, however, faces a variety of problems, such as for example rising postal and printing costs, which effect the cost of doing business. The success of a direct mailing is dependent on the number of responses created by the direct mailing (i.e., the response rate). As a result, blindly mailing a direct mail piece to everyone on a mailing list (e.g., mass mailings) can be costly and inefficient because the response rate will in all likelihood be low.

Cost-conscious direct marketers use their knowledge about persons identified on a mailing list (i.e., a prospect) to determine the best prospects to mail to. Usually, a marketer will use a set of descriptor variables about each prospect, such as for example demographics and credit card ownership, to target good prospects (i.e., prospects which will find the mailing interesting). For example, the Rao and Steckel model includes acquiring a set of descriptor variables and conducting a knowledge engineering session to screen the variables. In this regard, a marketing committee may be appointed and prior experience and intuition may be used to pick out the demographic variables most relevant to the response rate. After that, the probability of a response is modeled as a beta-logistic distribution and its parameters are estimated by maximum likelihood and a response score, $R(i)$ for each prospect i , and a profit score, $P(i)$ for each prospect i , are generated. Next, each prospect is assigned a value of $R(i) \times P(i)$, and each prospect is ranked high to low based on the assigned value of $R(i) \times P(i)$. The ranking of prospects is intended to account for both responsiveness and profitability of the direct mailing proposal.

There are at least two drawbacks to the Rao and Steckel model. First, the variable screening process based on opinion is subjective and error-prone. Second, using a simple distribution to describe the response probability in a high-dimensional (e.g., hundreds of attributes to a prospect), noisy environment (i.e., incomplete or missing data) is inadequate. In this regard, the drawback of using a simple distribution to describe the response probability is that a simple distribution assumes that the behaviors of various people will follow more or less a magic distribution and is governed by pure randomness. This assumption ignores the fact that there may be some reasons behind a person's response to a solicitation.

BRIEF SUMMARY OF THE INVENTION

Thus there is a particular need for a consistent and sustainable process for building response models which predict the likelihood of a prospect responding to a marketing solicitation. The present invention is a method for modeling market response rates. The method is used to evaluate and filter large contact lists with the aim of accomplishing two goals. The tactical goal is to improve a market response rate to cut costs associated with mailing, phone and electronic mail campaigns and produce more leads. The strategic goal is to assess the incremental risk of non-responsiveness associated with the incremental volume derived from growing a market in different directions (i.e., the tradeoffs between growing business (e.g., more solicitation) and risk of loss).

The method of the present invention uses an internal experience database and an external demographic database, coupled with variable screening schemes and non-parametric modeling techniques. The method comprises a number of steps. First, a data acquisition step associates descriptor variables with prospects. Second, a variable selection step identifies the descriptor variables in order to identify prospects most likely to respond to a direct marketing solicitation. Third, a model selection step examines and assesses a number of competitive algorithms, and selects the algorithm that will best predict the response rate. Fourth, a parameter estimation step ensures the best fit of data once an algorithm is chosen. Finally, a validation step ensures the robustness of the

modeling process. Robustness means the model will work even though the data in the future is likely to be different from the data used to build the model.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a flow diagram illustrating the steps of one embodiment of a method for modeling market response rates in accordance with the present invention;

5 Fig. 2 is a block diagram illustrating an example of cross-referencing of mailing records with multiple universal files in the data acquisition step of the method illustrated in Fig. 1;

Fig. 3 is a flow diagram illustrating steps included in the variable selection step of the method illustrated in Fig. 1;

10 Fig. 4 is a table showing an exemplary subset of Census Tract and Block Group variables selected in the variable selection step of the method illustrated in Fig. 1;

Fig. 5 is a table showing the results of testing of models constructed using training data;

15 Fig. 6 is a table showing the performance of ZIP5 classifiers;

Fig. 7 is a table showing the performance of Census Tract and Block Group classifiers;

Fig. 8 is a table showing the performance of Donnelley classifiers; and

Fig. 9 is a graph summarizing the results shown in Figs. 6-8.

DETAILED DESCRIPTION OF THE INVENTION

20 The first embodiment of the invention is a method for estimating response rates for a direct marketing campaign. For ease of explanation, this invention will be described with reference to a direct marketing campaign that uses the mail, however,

other processes can be used with this invention such as direct marketing with a phone, the Internet (electronic mail), fax machines, etc. A direct marketing campaign through the mail can provide a variety of information to the intended recipients. One possible example is to mail pieces advertising long term health care insurance. This invention should not be limited to advertising long term health care insurance and can be use for a variety of other insurance applications as well other areas that do not relate to insurance. The method of this invention associates demographic variables to prospects and uses non-parametric modeling techniques to predict mailing response rates for the prospects. The method is operable in two modes -- a training mode and a testing mode -- for cross-validation purposes. Specifically, the data set is divided into two sets, a training data set that is used to build the model, and a test data set that is used to test the robustness of the model. In the training mode, historical mailing data is analyzed off-line and a decision logic (i.e., a model) is formulated to estimate the mailing response rates. In the testing mode, the decision logic analyzes prospects on the fly and predicts the response rates for prospects.

Referring now to Fig. 1, the method first comprises the step of acquiring data. This step generally comprises attaching household or area level demographics to a prospect (e.g., a mailing record), randomly sampling the prospects, and splitting the randomly sampled prospects into a training set and a testing set. First, mailing records on a mailing list are cross referenced with a universal file (i.e., the entire data set) so that information regarding demographic variables associated with a mailing record are attached to the mailing record. Preferably, the mailing records are cross referenced with multiple universal files. For example, as shown in Fig. 2, the mailing records can be broken down into groups using universal files available from various vendors, including for example Donnelley, Census Tract and Block Group ("CTBG") and ZIP5. If multiple universal files are used, preferably the mailing records are broken down in to subgroups. Continuing the example, four groups of data can be created -- Group 1: Matched with Donnelley household demographic key; Group 2: Not matched with Donnelley, but matched with CTBG demographic key; Group 3: Not matched with either Donnelley or CTBG, but matched with ZIP5 demographic key; and Group 4: Not matched with

Donnelley, CTBG or ZIP5. Preferably, the mailing record is attached with individual, household and area level demographic information which are useful for identifying segments having the strongest relationship to the mailing response rate. Preferably, for each universe file used for cross referencing, equal number of responders and non-responders are included in the group. In this regard, all of the responders and a sub-sampling (i.e., random drawing of non-responders) of the non-responders are typically included because non-responders are much greater in size than that of responders. Each group is next randomly split into two sets - a training set and a testing set. For example, the training set may be about 2/3 of the size of the group, and the testing set may be about 1/3 of the group.

Referring again to Fig. 1, the method next comprises the step of variably selecting descriptor variables. Generally, it is desirable to use as few variables as possible in the presence of noise. This is often referred to as the "principle of parsimonious." There may be combinations (linear or nonlinear) of variables that are irrelevant to the underlying process, that due to noise in data appear to increase the prediction accuracy. Preferably, variables with the greater discrimination power in response prediction are selected. Generally, descriptor variables are selected using the misclassification rate as a measure of the discrimination power of each input variable given the same size of tree it constructs. In this regard, there are two types of misclassification, "wasted-mail" and "missed-opportunity." A model takes as input a list of prospects attached with demographic variables (Xs) and known response (Y's) and will produce as output four numbers -- first, the number of known responders being classified as responders (sensitivity of the classifier); second, the number of known responders being classified as non-responders (missed-opportunity); third, the number of known non-responders being classified as responders (wasted-mail); and fourth, the number of known non-responders being classified as non-responders (specificity of the classifier). It is preferred that both missed-opportunity and wasted-mail be minimized. However, tradeoffs may be necessary. For example, one performance evaluation criterion would be to minimize the misclassification cost, that is to minimize the sum of

misclassification rates weighted by cost: objective = minimize (# of wasted-mail x cost per wasted-mail + # of missed-opportunity x cost per missed-opportunity).

This step involves two parts. First, it is preferred that CART, a commercially available statistical algorithm for classification be used for variable selection. Assuming there are N input variables and one output variable, and that there are equal number of responders and non-responders in the training data set, a tree model is constructed for each input variable and the output variable is used to measure how good the variable is, as shown in Fig. 3. Next, the tree is allowed to grow until the size of each terminal node is preferably no smaller than 1/100 of the original data set. Next, the tree is pruned until the number of terminal nodes is preferably around 10, which provides a balance between robustness and accuracy. Next, the misclassification rate of the tree model is computed. (At this point, there are N tree models. Each tree has about 10 terminal nodes.) Next, each (N) tree model ranked in ascending order of their misclassification rates. Finally, the top 20 trees and their input variables are selected. For example, a subset of CTBG variables selected by CART are shown in Fig. 4. This step secondly involves selecting variables out of the available Donnelley, CTBG and ZIP5 variables. In this regard, each input variable is grouped into two samples: responders and non-responders. The mean difference of the two groups for the particular input variable is next tested. In addition, the variance difference of the two groups is tested. If both the mean and variance values are significant different, then the input variable is selected. The selection criteria in this case is that both P-values from two-sample T-test and F-test are significant at 0.01 level. The variables that are common to the two groups of variables are preferably used.

Referring again to Fig. 1, the method next comprises the step of selecting a classifier that will best model the mailing response rate. First, available classifiers that are to be considered as possible models are selected. For example, commercially available classifiers, including METROMAIL, Multivariate Adaptive Regression Splines, Logistic Regression, Neural Networks with Back-Propagation, CART and No Data Optimal Classifier (e.g., human intuition), can be selected. Next, the selected classifiers are compared using available universal files, such as for example, the ZIP5

universal file. In this regard, the selected classifiers are constructed using the selected universal file that is split into a training set, to construct the classifier, and a test set, to test the robustness of the constructed classifier. Preferably, multiple universal files are used. For example, the ZIP5 universal file is split into a training set and a testing set, and the training set is used to construct a number of different classifiers such as METROMAIL, Multivariate Adaptive Regression ("MARS"), Splines ("C4.5"), Logistic Regression ("LR"), Neural Networks with Back-Propagation ("NN-BP"), CART and No Data Optimal Classifier ("NDOC"). Next, the ZIP5 testing set is used to test the constructed classifiers for robustness. In this example, the ZIP5 universe file contained 8,407 responders and 454,732 non-responders.

With reference to Fig. 5, the results of the testing of each model constructed using the training data are shown using the test data to validate. This comprises the validation step shown in Fig. 1. In order to make a cost evaluation, the cost per missed-opportunity (i.e., no mailing was made to prospect which would have responded) is estimated to be \$17.85, while the cost per wasted-mail is estimated to be \$.33 (i.e., the cost of postage). Based on the results of the example, the CART classifier was determined to be the best. The results of the test, and the best classifier will vary according to the classifiers used, the universal file used, and the assumptions made.

Referring again to Fig. 1, the method further comprises the step of estimating the parameters. In this regard, it is noted that NODAC (no-data optimal classifier) is in essence a "brain-dead" approach. It does not utilize any data/information to reach its conclusion. It only depends on the gut feeling of prior probabilities and estimation of misclassification cost. Theoretically speaking, it assigns any observation to a class j that minimizes $\sum \pi(i)C(j|i)$, for all i , where $\pi(i)$ is the prior probability of class i and $C(j|i)$ is the cost of misclassifying class i as class j . Assuming class 1 to be responders, and class 0 to be non-responders, there are two costs—missed-opportunity, $C(0|1)$, and wasted-mail, $C(1|0)$. The latter is to misclassify class 0 as class 1, i.e., mistaken non-responders as responders. The former is vice versa, i.e., mistaken responders as non-responders. In this example, NODAC has only two choices. It either blindly mails out

to all prospects or does not mail at all. The decision is based on the minimum of the two numbers: $\pi(1)C(0|1)$ or $\pi(0)C(1|0)$. The latter was the total cost of wasted-mail (mail to all and see how much it costs): $\pi(0)C(1|0) = 0.98 \times \$0.33 = 0.3240$. The former was the total cost of missed-opportunity (do not mail at all): $\pi(1)C(0|1) = 0.018 \times \$17.85 = 0.3240$. As can be seen by this example, the two costs were almost the same. The cost estimates in this case are called the break-even costs for the prior estimates.

The parameters that can be set in tree structured classification include the priors, $\pi(i)$, and variable misclassification costs, $C(j|i)$, for class i and j . The priors are more or less fixed -- not much can be done about the 1.56% response rate on the average. Consequently, instead of bumping up the 1.56% response rate, it is preferred that the classifier's prediction accuracy is improved by using better estimates of misclassification costs.

As discussed above, break-even costs for the ZIP5 priors are \$17.85 and \$0.33 for missed-opportunity and wasted-mail, respectively. In this example, a high confidence on the estimate of wasted-mail is presumed to be within 10%. In contrast, the missed-opportunity depends on how the profit is modeled. Nevertheless, the lower bound of the figure is the break-even cost of missed-opportunity: \$17.85 for the ZIP5. In this regard, it is not worth doing any business if a lead's value is lower than that. On the other hand, if the estimate of missed-opportunity increased in the positive direction, we will be tempted to mail out to all prospects due to the fact that the cost of missed-opportunity would be too high. In other words, as the cost of missed-opportunity increases, the NODAC (no-data optimal classifier) will become more and more dominate.

By way of further examples, the following shows the relative performance of METROMAIL, NODAC and CART under three different cost estimates for the three demographics. Fig. 6 illustrates the performance of ZIP5 classifiers. Figs. 7 and 8 show the performance of CTBG and Donnelley, respectively.

There are two numbers in each cell of Fig 6. One is the total misclassification cost. The other is the percentage improvement over METROMAIL. If the missed-

opportunity cost is estimated as \$20, then CART is clearly the better classifier. If it is estimated as \$60, then both CART and NODAC give similar performance. If the estimate of missed-opportunity is further increased to \$100, then CART is acting exactly like NODAC (i.e., mailing out to all prospects). Note that the performance of NODAC is the same throughout the three different estimates of missed-opportunity because the break-even cost is around \$20. Consequently, NODAC just mails out to all prospects as long as the cost of missed opportunity is estimated greater than \$20.

A similar trend as ZIP5 is observed for CTBG, as shown in Fig. 7. CART is the better classifier if the missed-opportunity cost is around \$20. If the missed-opportunity is estimated beyond \$60, then both CART and NODAC behave in the same way. The CART's dominance over NODAC decreases as the missed-opportunity cost is increasing from \$20 to \$60.

From Fig. 8, it is clear that CART is the better classifier throughout various levels of cost estimates of missed-opportunity. CART has potential savings over METROMAIL by \$2.4 MM if the cost of per missed-opportunity is \$100.

The results shown in Figs. 6-8 are summarized in Fig. 9. The X-axis is the cost estimates of missed-opportunity, while the Y-axis is the dollars saved over METROMAIL classifier summed across the three demographics. If a lead's value is less than \$20, then it is not worth doing any business. Note that a lead's value is the same as the miss-opportunity cost. At the break-even cost, \$20, CART can save \$318,000 over the current METROMAIL classifier. If a lead's value is \$60, then either CART or NODAC can save up to \$1.8MM. If a lead is valued at \$100, then either CART or NODAC can save over \$4.4MM over the METROMAIL.

It is noted that CART is the better classifier if a lead is valued less than \$60. If a lead's value is greater than \$60, then there is not much gained using CART over NODAC. Assuming the missed-opportunity and the wasted-mail costs are \$100 and \$0.33, respectively, (i.e., the missed-opportunity cost is 303 times greater than that of the wasted-mail cost), it would break even if and only if at least one responder in 303

mailings was obtained. Knowing that the prior probability of response rate is 1.56%, the NODAC approach way would be used (i.e., mail out to all prospects).

5 It should be apparent that the method of the present invention provides a consistent and sustainable process for building response models which can be used in a variety of direct marketing scenarios that use the phone, Internet (electronic mail), fax machines, etc.

10 It is therefore apparent that there has been provided in accordance with the present invention, a method that fully satisfies the aims and advantages and objectives set forth herein. The invention has been described with reference to several embodiments, however, it will be appreciated that variations and modifications can be effected by a person of ordinary skill in the art without departing from the scope of the invention.

CLAIMS

1. A method for modeling market response rates comprising the steps of:

acquiring a list of prospects and attaching descriptor variables to said list of prospects;

variably selecting descriptor variables;

5 selecting a model by examining and assessing at least one algorithm; and

validating the model to ensure the robustness of the modeling process.

2. The method of claim 1, wherein said step of acquiring a list of prospects further comprises:

attaching demographics to a prospect;

randomly sampling the prospects; and

5 splitting the randomly sampled prospects into a training set and a testing set.

3. The method of claim 1, wherein said step of acquiring a list of prospects further comprises:

cross referencing said list of prospects with a universal file to cause variables associated with a mailing record to be attached to said prospect.

4. The method of claim 3, wherein said list of prospects is cross referenced with multiple universal files.

5. The method of claim 4, wherein said list of prospects are cross referenced with at least two of the group of universal files consisting of: Donnelley, Census Tract and Block Group and ZIP5.

6. The method of claim 1, wherein said step of acquiring a list of prospects further comprises selecting an equal number of responders and non-responders for each universe file used for cross referencing.

7. The method of claim 1, wherein said step of variably selecting descriptor variables further comprises using the misclassification rate as a measure of the discrimination power of each input variable.

8. The method of claim 1, wherein said step of variably selecting descriptor variables further comprises:

selecting CART, and using CART to performing the steps of:

constructing a tree model for each input variable available for selection;

5 growing the tree until the size of each terminal node is preferably no smaller than 1/100 of the original data set;

pruning the tree until the number of terminal nodes reaches a predetermined number;

computing the misclassification rate of each tree model;

10 ranking each tree model in ascending order of their misclassification rates; and

selecting at least one input variable based on said ranking.

9. The method of claim 1, wherein said step of variably selecting descriptor variables further comprises:

selecting at least one variable out of the available Donnelley, Census Tract and Block Group and ZIP5 variables;

5 grouping each input variable into a responder group and a nonresponder group;

determining the mean difference of the responder group and nonresponder group for the input variable selected;

determining the variance difference of the responder group and nonresponder group for the input variable selected;

10 selecting the input variable if both the mean and variance values are significant different.

10. The method of claim 1, wherein said step of selecting a model by examining and assessing at least one algorithm further comprises:

selecting multiple classifiers for examination and testing; and

comparing the selected classifiers using a universal file.

11. The method of claim 10, wherein said step of selecting an algorithm comprises selecting an algorithm from group consisting of METROMAIL, Multivariate Adaptive Regression Splines, Logistic Regression, Neural Networks with Back-Propagation, CART and No Data Optimal Classifier.

12. The method of claim 10, wherein said universal file is the ZIP5 universal file.

13. The method of claim 10, wherein said step of comparing the selected classifiers using a universal file further comprises:

splitting the universal file that is split into a training set and a testing set; and

constructing the classifier using the training set.

14. The method of claim 13, where said step of validating the model to ensure the robustness of the modeling process further comprises:

testing each classifier using the testing set.

15. The method of claim 1, further comprising the step of estimating the parameters to ensure that the best fit of data once an algorithm is chosen.

16. The method of claim 1, wherein said method is applied to modeling response rates in the insurance sector.

17. A method for modeling market response rates comprising the steps of:

acquiring a list of prospects;

attaching demographics to said prospect;

randomly sampling the prospects;

5 splitting the randomly sampled prospects into a training set and a testing set;

variably selecting descriptor variables using the misclassification rate;

selecting multiple classifiers;

selecting a universal file;

comparing the selected classifiers using the selected universal file by constructing

10 the classifier using said training set;

validating the selected classifier by testing each classifier using the testing set.

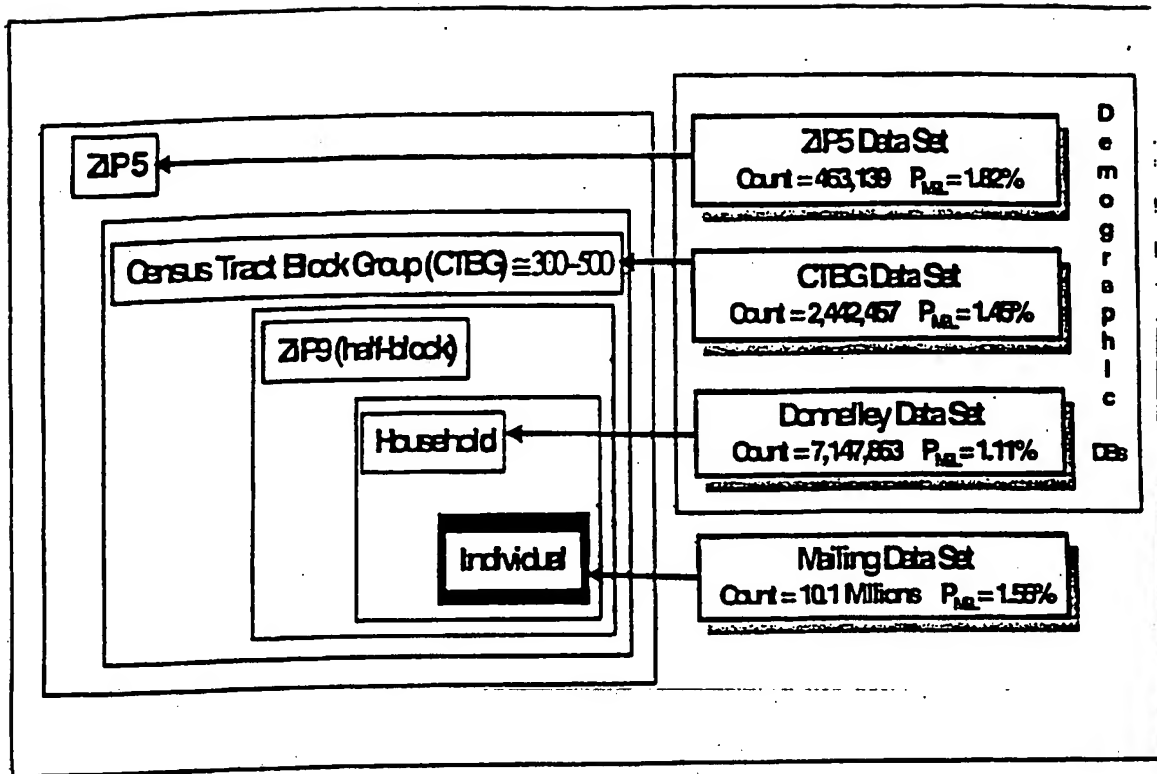


Fig. 2

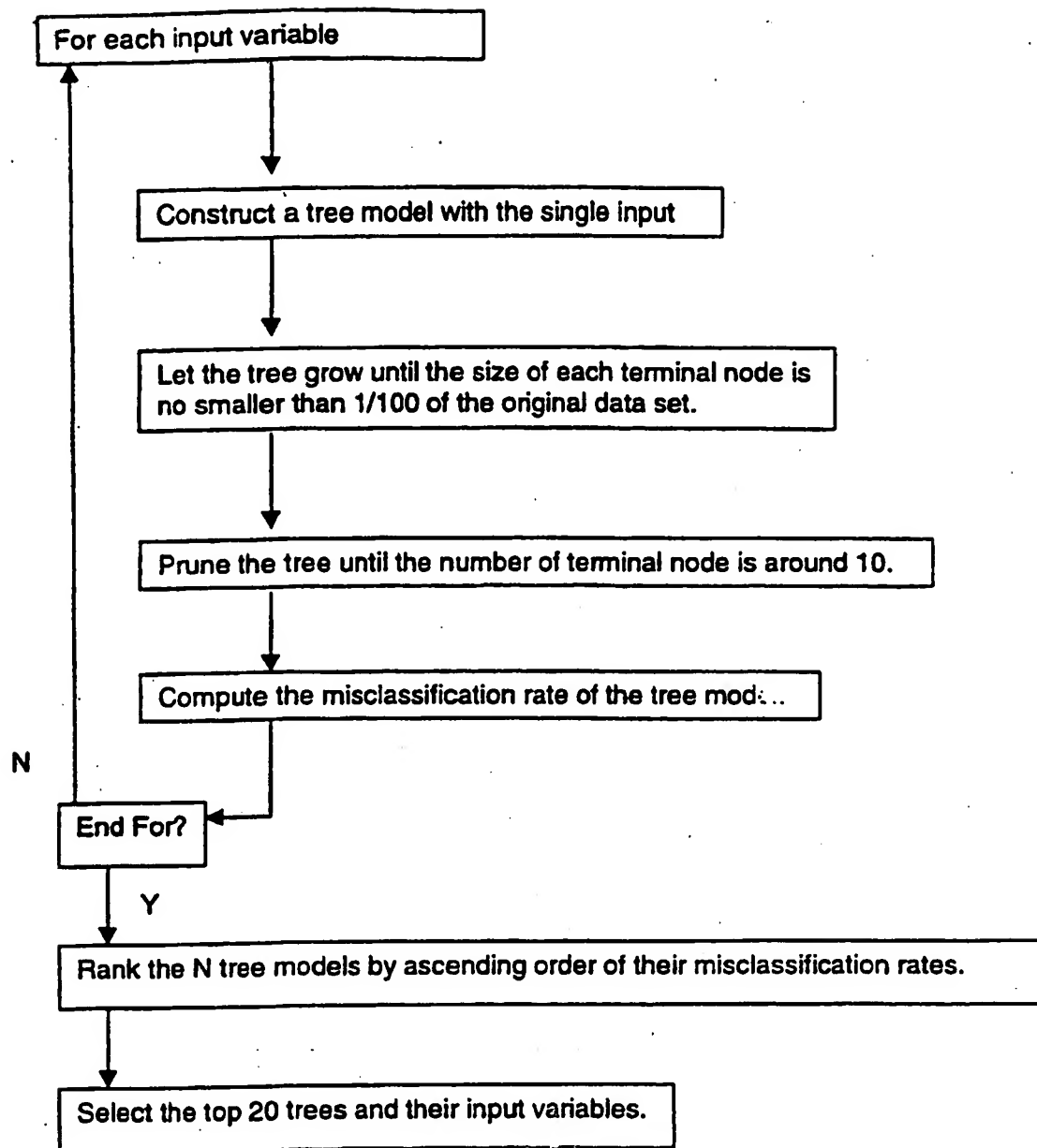


Fig. 3

	Variable Description
1	Easy weather index
2	Per capital income
3	Easy quality of life index
4	Median value of owner occupied houses
5	Population in 1997 in the CTBG
6	Restaurant index
7	Car, truck, van used to work
8	Urban population in the CTBG
9	Culture index

Fig. 4

	# of missed- opportunit y	# of wasted- mail	Misclassificat i n cost	% improved over METROMAIL
METROMAIL	2,480	343,323	\$157,566	0.0%
C4.5	2,430	285,572	\$137,607	12.7%
MARS	3,918	203,265	\$137,008	13.0%
LR	4,700	172,798	\$140,910	10.6%
NN-BP	8,071	13,187	\$148,414	5.8%
CART	4,802	145,983	\$133,890	15.0%
NODAC	0	454,732	\$150,062	4.8%

Fig. 5

	\$20	\$60	\$100
METROMAIL	\$157,566 (0.0%)	\$247,220 (0.0%)	\$378,911 (0.0%)
NODAC	\$150,062 (4.8%)	\$150,062 (39.3%)	\$150,062 (60.4%)
CART	\$133,890 (15.0%)	\$151,379 (38.8%)	\$150,062 (60.4%)

Fig. 6

	\$20	\$60	\$100
METROMAIL	\$818,586 (0.0%)	\$1,315,204 (0.0%)	\$2,060,185 (0.0%)
NODAC	\$794,583 (2.9%)	\$794,583 (39.6%)	\$794,583 (61.4%)
CART	\$721,913 (11.8%)	\$794,583 (39.6%)	\$794,583 (61.4%)

Fig. 7

	\$20	\$60	\$100
METROMAIL	\$2,309,073 (0.0%)	\$3,499,839 (0.0%)	\$5,286,130 (0.0%)
NODAC	\$2,309,455 (0.0%)	\$2,309,455 (34.0%)	\$2,976,675 (77.6%)
CART	\$2,109,965 (8.6%)	\$2,309,455 (34.0%)	\$2,976,675 (77.6%)

Fig. 8

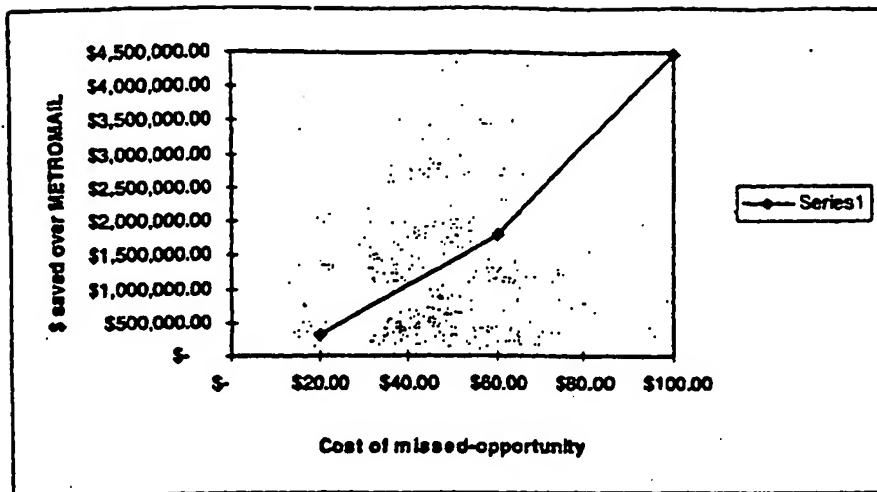


Fig. 9